Available online at www.sciencedirect.com
SCIENCE DIRECT
Applied Mathematics Letters 17 (2004) 683–689
Applied Mathematics Letters
www.elsevier.com/locate/aml

An Objective Penalty Function Method for Nonlinear Programming

ZHIQING MENG
College of Business and Administration, Zhejiang University of Technology
Hangzhou, Zhejiang 310032, P.R. China
mengzhiqing@163.net

QIYING HU
College of International Business and Management, Shanghai University
Shanghai 201800, P.R. China

CHUANGYIN DANG
Department of Manufacturing Engineering and Engineering Management
City University of Hong Kong
Kowloon, Hong Kong
mecdang@cityu.edu.hk

XIAOQI YANG
Department of Applied Mathematics, The Hong Kong Polytechnic University
Hung Hom, Hong Kong
mayangxq@polyu.edu.hk

(Received August 2002; revised and accepted August 2003)

Abstract—In this paper, we propose a novel objective penalty function for inequality constrained optimization problems. The objective penalty function differs from any existing penalty function and also has two desired features: exactness and smoothness if the constraints and objective function are differentiable. An exact penalty result is proved for the objective penalty function. In addition to these results, based on the objective penalty function, we develop an algorithm for solving the original problem and show its convergence under some mild conditions.

Keywords—Nonlinear programming, Exact penalty function, Objective penalty function.

1. INTRODUCTION

The problem we consider in this paper is as follows:

$$\begin{aligned} \min \quad & f_0(x), \\ \text{s.t.} \quad & f_i(x) \le 0, \quad i \in I = \{1, 2, \dots, m\}, \end{aligned} \tag{P}$$

The authors would like to thank anonymous referees' comments and remarks that help us to improve the presentation of this paper considerably. This research was partially supported by SRG 7001150 and the National Natural Science Foundation of China under Grant 70271021.

0893-9659/04/$ - see front matter Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX
doi:10.1016/j.aml.2003.08.010

where $f_i : R^n \to R \cup \{+\infty\}$, $i \in I_0 = \{0, 1, 2, \ldots, m\}$. Let $X = \{x \in R^n \mid f_i(x) \le 0,\ i \in I\}$, which is the feasible set of (P).

The penalty function method is an important approach to solving (P). To obtain a solution of (P), the penalty function method solves a sequence of unconstrained optimization problems. In recent years, researchers have been focusing on theory and practical applications of penalty functions. In many studies, a penalty function of (P) is defined as

$$F(x,\rho) = f_0(x) + \rho \sum_{i \in I} \max\{f_i(x), 0\}^2,$$

and the penalty optimization problem of (P) is defined as

$$F(x,\rho), \qquad \text{s.t. } x \in R^n. \tag{P$_\rho$}$$

The penalty function $F(x,\rho)$ is smooth (or differentiable) if the constraints and objective function are differentiable, but not exact. The penalty function $F(x,\rho)$ is exact if there is some ρ^* such that an optimal solution to (P$_\rho$) is also an optimal solution to (P) for any given $\rho \ge \rho^*$. In 1967, Zangwill [1] proposed an exact penalty function

$$F_1(x,\rho) = f_0(x) + \rho \sum_{i \in I} \max\{f_i(x), 0\},$$

and the penalty optimization problem of (P) is defined as

$$F_1(x,\rho), \qquad \text{s.t. } x \in R^n. \tag{EP$_\rho$}$$

The penalty function $F_1(x,\rho)$ is exact, but not smooth. The study of exact penalty functions attracts many researchers' interests. Han and Mangasarian (cf. [2]) presented another exact penalty function for nonlinear programming. Rosenberg (cf. [3]) gave a globally convergent algorithm for convex programming based on an exact penalty function. Lasserre (cf. [4]) proposed a globally convergent algorithm for some exact penalty functions. Dippillo and Grippo (cf. [5]) developed an exact penalty function method with global convergence properties for nonlinear programming problems. Zenios *et al.* (cf. [6]) found a smooth penalty function algorithm for network-structured problems. Pinar and Zenios (cf. [7]) studied smooth exact penalty functions for convex constrained optimization problems. Mongeau and Sartenaer (cf. [8]) analyzed automatic decrease of the penalty parameter in exact penalty function methods.

Recently, Rubinov *et al.* (cf. [9,10]), Yang and Huang (cf. [11]), and Huang and Yang (cf. [12]) studied a new penalty function called a nonlinear Lagrangian penalty function, which is defined as

$$F_k(x,\rho) = \left[f_0(x)^k + \rho \sum_{i \in I} \max\{f_i(x), 0\}^k \right]^{1/k},$$

where $k > 0$, and the penalty optimization problem of (P) is defined as

$$F_k(x,\rho), \qquad \text{s.t. } x \in R^n. \tag{EPk$_\rho$}$$

When $k = 1$, problem (EPk$_\rho$) is the same as (EP$_\rho$). So, $F_k(x,\rho)$ can be considered as a generalization of $F_1(x,\rho)$. The penalty function $F_k(x,\rho)$ is smooth for $k > 1$ if the constraints and objective function are differentiable, but not smooth for $0 < k \le 1$. Although an exact penalty result was obtained under some conditions in [9,10], it is not easy to check the conditions. On the other hand, almost all the results in [9–12] have a necessary assumption that the objective function $f_0(x)$ is positive on X. Therefore, exactness and smoothness of the penalty function $F_k(x,\rho)$ for $k \ne 1$ are not better than those of $F(x,\rho)$ or $F_1(x,\rho)$.

All penalty function algorithms need to increase penalty parameter ρ sequentially in order to solve (P). So does the exact penalty function because we often do not know exactly how big the penalty parameter ρ is. But, in practical computing, it is impossible to take too big a penalty parameter ρ. Hence, exactness and smoothness become a key concern for a penalty function. It is important to find out better conditions for an exact penalty function. Consequently, we have to look for other types of penalty functions that give us new prospects to solve the problems the existing exact and smooth penalty functions have.

In this paper, for a given number $p > 1$, we study an objective penalty function

$$F(x, M) = (f_0(x) - M)^2 + \sum_{i \in I} f_i^+(x)^p,$$

where $M \in R$ is called an objective penalty parameter and

$$f_i^+(x) = \max\{0, f_i(x)\}, \qquad i \in I.$$

If f_i ($\forall\, i \in I_0$) is first-order differentiable at any $x \in R^n$, then $f_i^+(x)^p$ is first-order differentiable at any $x \in R^n$, and so is $F(x, M)$.

The rest of this paper is organized as follows. In Section 2, we prove an exact penalty result of the objective penalty function $F(x, M)$ and give an algorithm to solve the original problem (P), which converges without convex condition and provides an alternative way to solve constrained optimization problems. We conclude this paper with some remarks in Section 3.

2. AN OBJECTIVE PENALTY FUNCTION METHOD

Consider the following nonlinear unconstrained optimization problem:

$$\min F(x, M), \qquad \text{s.t. } x \in Y, \tag{P(M)}$$

where $Y \supset X = \{x \in R^n \mid f_i(x) \leq 0,\ i \in I\}$. If an optimal solution to (P(M)) for some M is an optimal solution to (P), then M is called an exact penalty parameter. Next, we give a sufficient condition that the penalty function $F(x, M)$ is exact for (P).

THEOREM 2.1. *Let x^* be an optimal solution to (P). For some M, let x_M^* be an optimal solution to (P(M)) and a feasible solution to (P) with $F(x_M^*, M) > 0$. If $M \leq f_0(x^*)$, x_M^* is an optimal solution to (P).*

PROOF. By the assumption, we have

$$(f_0(x_M^*) - M)^2 \leq (f_0(x) - M)^2, \qquad \forall\, x \in X. \tag{1}$$

It follows from $M \leq f_0(x^*)$ that

$$f_0(x_M^*) - M \geq f_0(x_M^*) - f_0(x^*) \geq 0, \tag{2}$$

since x_M^* is a feasible solution and x^* is an optimal solution to (P). Suppose that there is some $x \in X$ such that $f_0(x) < M$. Then $f_0(x) < M \leq f_0(x^*)$, which contradicts the fact that x^* is optimal to (P). Therefore, for any $x \in X$, we have $f_0(x) - M \geq 0$, which together with (1) and (2) imply that

$$f_0(x_M^*) - M \leq f_0(x) - M, \qquad \forall\, x \in X.$$

So x_M^* is an optimal solution to (P). This completes the proofs. ∎

THEOREM 2.2. *Let the feasible set X be connected and compact, f_0 a continuous function in R^n, $M_* = \min_{x\in X} f_0(x)$, and $M^* = \max_{x\in X} f_0(x)$. For some M, let x_M^* be an optimal solution to (P(M)). Then*

(i) $M_* \le M \le M^*$ *if* $F(x_M^*, M) = 0$.
(ii) $M \le M_*$ *if* $F(x_M^*, M) > 0$ *and* $M \le M^*$. *Furthermore,* x_M^* *is an optimal solution to (P) if* x_M^* *is a feasible solution to (P).*

PROOF. The result of (i) is obvious.

(ii) If $M_* \le M$, we have $M_* \le M \le M^*$. Since f_0 is continuous on the connected and compact set X, there is some $x_0 \in X$ such that $M = f_0(x_0)$. So, we get $F(x_0, M) = 0$. On the other hand, since x_M^* is optimal to (P(M)) and $F(x_M^*, M) > 0$, hence, $0 < F(x_M^*, M) \le F(x_0, M) = 0$, which leads to a contradiction. Therefore, $M \le M_*$. If x_M^* is a feasible solution to (P), we deduce that x_M^* is an optimal solution to (P) from Theorem 2.1. This completes the proofs. ∎

We remark that the result of Theorem 2 is better than ones of the existing exact penalty functions.

According to Theorem 2.2, we develop an algorithm to generate a globally optimal solution to (P), which is called an OPFM Algorithm.

AN OPFM ALGORITHM.

Step 1: Choose $p > 1$, $\epsilon > 0$, $x^0 \in X$, and a_1 satisfying $a_1 < f_0(x^0)$. Let $k = 1$, $b_1 = f(x^0)$, and $M_1 = (a_1 + b_1)/2$, and go to Step 2.
Step 2: Solve $\min_{x\in Y} F(x, M_k)$. Let x^k be the optimal solution obtained.
Step 3: If x^k is not a feasible solution to (P), let $b_{k+1} = b_k$, $a_{k+1} = M_k$, $M_{k+1} = (a_{k+1} + b_{k+1})/2$, and go to Step 5. Otherwise, $x^k \in X$, and go to Step 4.
Step 4: If $F(x^k, M_k) = 0$, let $a_{k+1} = a_k$, $b_{k+1} = M_k$, and $M_{k+1} = (a_{k+1} + b_{k+1})/2$, and go to Step 5. Otherwise, $F(x^k, M_k) > 0$, x^k is an optimal solution to (P), and the algorithm terminates.
Step 5: If $|b_{k+1} - a_{k+1}| < \epsilon$, then the algorithm terminates and x^k is an approximately optimal solution to (P). Otherwise, let $k = k + 1$, and go to Step 2.

In the following theorem, we show that under some mild conditions, the OPFM Algorithm converges.

THEOREM 2.3. *Let X be connected and compact. Let $a_1 < M_* = \min_{x\in X} f_0(x)$. Assume that f_i ($i \in I_0$) are continuous in R^n and that the level set $L(\alpha, \beta) = \{x \mid \alpha \le f_0(x) \le \beta\}$ is bounded for any given $\alpha, \beta \in R$. In the OPFM Algorithm, let $\epsilon = 0$, and $\{x^k\}$ be a sequence generated by the OPFM Algorithm.*

(i) *If the sequence $\{x^k\}$ is finite (or the OPFM Algorithm terminates at the k^{th} iterations), then x^k is an optimal solution to (P).*
(ii) *If the sequence $\{x^k\}$ is infinite, then $\{x^k\}$ is bounded and any limit point of it is an optimal solution to (P).*

PROOF. We first show that $\{a_k\}$ increases and $\{b_k\}$ decreases with

$$a_k < M_k < b_k, \qquad k = 1, 2, \dots, \tag{3}$$

and

$$b_{k+1} - a_{k+1} = \frac{b_k - a_k}{2}, \qquad k = 1, 2, \dots. \tag{4}$$

We use the induction method to prove it. It is clear that $a_1 < M_1 < b_1$ and $b_2 - a_2 = (b_1 - a_1)/2$ by the OPFM Algorithm. Suppose that (3) and (4) hold for some $k > 1$. For $k + 1$, in Step 3, we have $b_{k+1} = b_k$ and $a_{k+1} = M_k$, which results in that

$$M_{k+1} = \frac{a_{k+1} + b_{k+1}}{2} < \frac{b_k + b_k}{2} = b_k = b_{k+1}$$

and

$$M_{k+1} = \frac{a_{k+1} + b_{k+1}}{2} > M_k = a_{k+1}.$$

We obtain $b_k = b_{k+1}$, $a_k < a_{k+1}$, and $a_{k+1} < M_{k+1} < b_{k+1}$, which implies

$$b_{k+1} - a_{k+1} = b_k - M_k = \frac{b_k - a_k}{2}.$$

On the other hand, in Step 4, we have $a_{k+1} = a_k$ and $b_{k+1} = M_k$. So, we get

$$M_{k+1} = \frac{a_{k+1} + b_{k+1}}{2} < \frac{M_k + M_k}{2} = b_{k+1}$$

and

$$M_{k+1} = \frac{a_{k+1} + b_{k+1}}{2} > \frac{a_k + a_k}{2} = a_k = a_{k+1}.$$

We also obtain $a_k = a_{k+1}$, $b_k > b_{k+1}$, and $a_{k+1} < M_{k+1} < b_{k+1}$, which implies

$$b_{k+1} - a_{k+1} = M_k - a_k = \frac{b_k - a_k}{2}.$$

Hence, (3) and (4) hold and $\{a_k\}$ and $\{b_k\}$ converge. Let $a_k \to a^*$ and $b_k \to b^*$. By (4), we have $a^* = b^*$. Therefore, $\{M_k\}$ converges to a^*.

(i) For $\epsilon = 0$, by (4), we conclude that the stopping criterion $|b_{k+1} - a_{k+1}| < \epsilon$ will not occur. Hence, when the OPFM Algorithm terminates at the k^{th} iteration, it must terminate at Step 4, x^k is a feasible solution to (P), and $F(x^k, M_k) > 0$. By Theorem 2.2, x^k is an optimal solution to (P).

(ii) Since x^k is an optimal solution to $\min_{x\in Y} F(x, M_k)$, we have

$$F\left(x^k, M_k\right) \le \left(f_0\left(x^0\right) - M_k\right)^2, \qquad k = 1, 2, \ldots.$$

From $\lim_{k\to\infty} M_k = a^*$, we know that there is some $L > 0$ such that

$$LF\left(x^k, M_k\right)^2, \qquad k = 1, 2, \ldots.$$

Then

$$LF\left(x^k, M_k\right)^2 \ge \left(f_0\left(x^k\right) - M_k\right)^2, \qquad k = 1, 2, \ldots.$$

By (3), we have

$$a_1 - L^{1/2} \le M_k - L^{1/2} \le f_0\left(x^k\right) \le M_k + L^{1/2} < b_1 + L^{1/2}, \qquad k = 1, 2, \ldots.$$

So, $\{f_0(x^k)\} \subset (a_1 - L^{1/2}, b_1 + L^{1/2})$. By the assumption, $\{x^k\}$ is bounded.

Now, we show the second conclusion. Without loss of generality, let $x^k \to x^*$. We have shown that

$$a_k < M_k < b_k, \qquad k = 1, 2, \ldots$$

and $\{a_k\}$, $\{b_k\}$, and $\{M_k\}$ converge to a^*. By Step 3 and Theorem 2.2(ii), we know $a_k < M_*$, $k = 1, 2, \ldots$. By Step 4 and Theorem 2.2(i), we know $M_* \le b_k$, $k = 1, 2, \ldots$. Hence, letting $k \to +\infty$, we obtain $a^* = M_*$ since f_i $(i \in I_0)$ are continuous in R^n. Let y^* be an optimal solution to (P). Then $M_* = f_0(y^*)$. From $F(x^k, M_k) \le F(y^*, M_k) = (f_0(y^*) - M_k)^2$ and letting $k \to +\infty$, we obtain

$$F(x^*, M_*) \le 0,$$

which implies $M_* = f_0(x^*)$. Therefore, x^* is an optimal solution to (P). This completes the proofs. ∎

When $\epsilon > 0$, the OPFM Algorithm terminates within a finite number of iterations by Theorem 2.3. Thus, the OPFM Algorithm provides an alternative method to solve (P). In the OPFM Algorithm, we do not need to increase the penalty parameter M to ∞, which differs from those other penalty function methods in [1, 2, ..., 12].

EXAMPLE 2.1. Consider

$$\begin{aligned} \min \quad & x_1 + x_2, \\ \text{s.t.} \quad & x_1^2 - x_2 \le 0, \quad 0 \le x_1 \le 10. \end{aligned} \tag{PP}$$

The optimal solution of (PP) is $(x_1^*, x_2^*) = (0, 0)$ and the optimal objective value is 0. The objective penalty function is defined by

$$F(x, M) = (x_1 + x_2 - M)^2 + \left(\max\left\{0, x_1^2 - x_2\right\}^p + \max\{0, -x_1\}^p + \max\{0, x_1 - 10\}^p\right).$$

Let $Y = \{(x_1, x_2) \mid 0 \le x_1 \le 100,\ 0 \le x_2 \le 100\}$. We apply the OPFM Algorithm to solve (PP) on Matlab 6.1.

(1) Let $p = 4$, $x^0 = (2, 4) \in X$, $a_1 = -4$, $b_1 = 6$, $M_1 = 1$. We have an optimal solution $x^1 = (0.3333, 0.6667)$ to $\min_{x \in Y} F(x, 1)$. Since $F(x^1, 1) = 0$, we get $a_2 = -4$, $b_2 = 1$, $M_2 = -1.5$.

(2) Again, we solve $\min_{x \in Y} F(x, -1.5)$ and obtain its optimal solution $x^2 = (0, 0)$. Since $F(x^2, -1.5) = 2.25 > 0$, $x^2 = (0, 0)$ is an optimal solution to (PP) by Theorem 2.1.

3. CONCLUSIONS

In this paper, we have presented a novel objective penalty function with an objective parameter. We have shown that the objective penalty function is exact under some conditions. We have also developed an OPFM Algorithm to solve (P) based on the objective penalty function and proved its global convergence. The OPFM Algorithm differs from existing penalty function algorithms. It possesses good convergence property and provides another appealing approach for us to study (P) further.

REFERENCES

1. W.I. Zangwill, Nonlinear programming via penalty function, *Management Science* **13**, 334–358, (1967).
2. S.P. Han and O.L. Mangasarian, Exact penalty function in nonlinear programming, *Mathematical Programming* **17**, 251–269, (1979).
3. E. Rosenberg, Globally convergent algorithms for convex programming, *Mathematics of Operational Research* **6**, 437–443, (1981).
4. J.B. Lasserre, A globally convergent algorithm for exact penalty functions, *European Journal of Operational Research* **7**, 389–395, (1981).
5. G. Dippillo and L. Grippo, An exact penalty function method with global convergence properties for nonlinear programming problems, *Mathematical Programming* **36**, 1–18, (1986).
6. S.A. Zenios, M.C. Pinar and R.S. Dembo, A smooth penalty function algorithm for network-structured problems, *European J. of Operational Research* **64**, 258–277, (1993).
7. M.C. Pinar and S.A. Zenios, On smoothing exact penalty functions for convex constraints optimization, *SIAM Journal on Optimization* **4**, 486–511, (1994).
8. M. Mongeau and A. Sartenaer, Automatic decrease of the penalty parameter in exact penalty functions methods, *European Journal of Operational Research* **83**, 686–699, (1995).
9. A.M. Rubinov, B.M. Glover and X.Q. Yang, Extended Lagrange and penalty functions in continuous optimization, *Optimization* **46**, 327–351, (1999).
10. A.M. Rubinov, B.M. Glover and X.Q. Yang, Decreasing functions with applications to penalization, *SIAM Journal on Optimization* **10**, 289–313, (1999).

11. X.Q. Yang and X.X. Huang, A nonlinear Lagrangian approach to constrained optimization problems, *SIAM Journal on Optimization* **11**, 1119–1144, (2001).
12. X.X. Huang and X.Q. Yang, Duality and exact penalization for vector optimization via augmented Lagrangian, *Journal of Optimization Theory and Applications* **111**, 615–640, (2001).